Available online at www.sciencedirect.com

SciVerse ScienceDirect

Int. J. Human-Computer Studies 70 (2012) 552–560

International Journal of Human-Computer Studies

www.elsevier.com/locate/ijhcs

Detecting system failures from durations and binary cues

Nir Shahar, Joachim Meyer*, Michael Hildebrandt, Vered Rafaely

Department of Industrial Engineering and Management, Ben - Gurion University of the Negev, Beer - Sheva, Israel

Received 17 May 2011; received in revised form 11 February 2012; accepted 13 March 2012
Communicated by G. Lindgaard
Available online 23 March 2012

Abstract

Durations are often used to judge the status of an invisible process. However, the apparent duration of an interval depends on the actual duration and on other variables, such as the workload during the interval and the person's expectations. An experiment dealt with the use of durations as an information source on the state of an invisible process and the effects of expectations and workload on decisions regarding the process. Eighty-nine participants observed a computerized simulation of a process which could be either intact or faulty, with intact processes ending on average sooner than faulty ones, and they had to indicate whether or not the process is intact and to estimate its duration. A binary cue with either intermediate or no validity indicated whether the process was supposedly intact or not, generating expectations about the duration of the process. Perceived durations and the decisions about the intactness of a process depended on the actual process duration, as well as on the expectations generated by the binary cue. In addition, task workload affected time estimates, but it had no effect on participants' tendency to adhere to cue recommendations or their ability to distinguish between intact and faulty processes. Results show that users' duration-based decisions about the status of a computerized process are affected by internal and external cues. While users can use durations as an information source, they should, whenever possible, be accompanied by additional indicators, lowering the inherent uncertainty in the duration estimation process.

Keywords: Failure detection; Signal detection theory; Time estimation; Apparent duration; Categorization decisions; Decision making

1. Introduction

Users often wait for a system to complete some process, for instance when saving or accessing files, sending emails, or establishing the connection to a network. These processes usually complete successfully. However, they do occasionally fail, and then the user has to restart the process or has to take some corrective action. Users must decide whether or not the process progresses as expected to decide on the appropriate response. This is often difficult, because many processes are invisible to the user, and users face uncertainty regarding the process outcome. Waiting time is an important information source (e.g., Shaefer, 1990), and unusual event durations, i.e., events that are too long or too short, often indicate failures in the process.

Time is unique as an information source in that it is not directly observable. Instead, without a stop watch or another timing device, we assess time by evaluating a cumulative subjective experience. This experience is not only affected by the actual passing of time, but also by other variables, which may affect decisions that are based on the apparent duration of events. This study concerns the influence of a number of variables on perceived durations and on binary decisions that are based on durations.

When people wait for a process to end and observe the time it takes, they are engaged in prospective time estimation, in which the individual knows in advance that he or she will need to estimate the duration of an event (Zakay and Block, 1997). Perceived duration in prospective time estimation is affected by expectations and workload (attention). Jones and Boltz (1989) proposed the *expectancy contrast model*, according to which events differ in the extent to which their structure is coherent. At one end of the continuum are events with structured beginnings and ends (rather than random ones), and their structure has time characteristics that generate expectations for completion time. Examples are progress bars or percentage numbers in computer programs

*Corresponding author. Tel.: +972 8 6472216; fax: +972 8 6472958.
E-mail address: Joachim@bgu.ac.il (J. Meyer).

1071-5819/$ - see front matter
http://dx.doi.org/10.1016/j.ijhcs.2012.03.001

that show the gradual completion of some process. At the other end are incoherent events without time cues, or events in which time cues provide no information about the expected duration of the event (e.g., an hourglass icon on the screen). Users generally prefer structured displays to unstructured ones, although the two displays will not necessarily lead to different apparent durations (Meyer et al., 1995, 1996).

Jones and Boltz (1989) claim that structured events generate future oriented attending (i.e., expectations) that influences time estimation. Events that end within the expected time frame induce relatively accurate time estimates; events that exceed the expected time frame generate overestimation of durations, and events that end before the expected time frame lead to underestimation of durations.

A person's activity during the time interval also affects the prospective time estimates, with activities that generate high workload reducing apparent durations (Zakay, 1989; Zakay and Block, 1997; Zakay et al., 1999; and see Block et al., 2010 for a recent meta-analysis and review of studies on duration estimates and workload). The effect of workload is typically studied in a dual-task paradigm, in which participants simultaneously perform a primary task and a secondary task. Performing a secondary task while attending to the passage of time (i.e., the primary task) disrupts time estimation (Brown, 1985, 1997), with people typically underestimating durations when performing dual tasks (e.g., Brown, 1997). The effect of the secondary task on time perception has been explained in terms of *resource allocation theory*. This theory posits that attentional resources are limited. When performing a dual-task, the two tasks compete for attentional resources. Allocation of resources to the secondary task limits the resources available for time estimation. Therefore participants consider only part of the time information, leading to underestimation of time (Brown, 1997). Taatgen et al. (2007) developed a cognitive-architecture model of the process, using the ACT-R architecture to model the effects of cognitive efforts on time perception.

Users frequently judge the status of a process by its duration (did events occur too late or did one have to wait too long for something to happen?). These are usually simple binary decisions whether a process is intact or faulty, and the user has only two possible response options—to act as if the process is intact or to act as if it is faulty. Examples for such temporal decisions are abundant. For instance, software applications, such as email or search engines, often have predictable response times. When an application responds uncharacteristically fast or slow, the user may infer that there is a problem with the underlying process (e.g., the database server or the network connection), and take appropriate action such as restarting the application. Another domain where temporal judgments are frequently needed is process control. Many components in industrial plants have time-lagged responses, and it is often important that operators learn about these response characteristics. For instance, coolant pumps in nuclear power plants require between several seconds and half a minute to reach nominal flow rates. The control room operator is expected to monitor the start-up of pumps until the nominal flow rate is reached. Temporal information, i.e., the pump's expected spool-up time, can be used to detect potential problems with the system. Skraaning and Nihlwing (2008) demonstrated that explicit graphical representation of temporal information can improve the performance of nuclear power plant control room operators for challenging, knowledge-based simulator scenarios.

Binary decisions, like the ones whether a process is intact of faulty, can be analyzed with *Signal Detection Theory* (SDT; Green and Swets, 1966; Macmillan and Creelman, 2005). It quantifies distinctions between signal (i.e., a faulty process) and noise events (i.e., an intact process). The decision maker's ability to distinguish between signal and noise is referred to as sensitivity (d'). The tendency to detect signal or noise (i.e., the bias) is quantified through the decision criterion (β), which represents the value or the intensity of a stimulus above which the decision maker would identify it a signal. It is affected by the decision maker's prior expectations about the occurrence of a signal and the values of the four possible decision outcomes, commonly referred to as Hit (a signal is correctly identified as a signal; e.g., a fault is correctly detected), False Alarm (FA) (a noise event is incorrectly identified as a signal; e.g., a fault is declared when there is none), Miss (a signal is incorrectly identified as noise; e.g., the system is considered to be intact when there actually is a fault), and Correct Rejection (CR) (a noise event is correctly identified as noise; e.g., the system is considered to be intact when it is indeed intact). The decision maker should adopt a liberal decision criterion, in which both Hit and FA outcomes are more likely, when a signal is probable or when the expected reward for Hit or the expected loss for Miss are high. In contrast, users should adopt a conservative criterion, in which probabilities for CR and Miss outcomes are high, when signals are unlikely or when the expected reward for CR or the expected loss for FA are high.

Users often rely on more than one information source for their decisions. For instance, in many systems alarms, alerts, or other indicators provide indications, and this information is combined with information from other sources. In the experiment we report here we look at the effect of a simple binary cue that indicates whether a process is intact or faulty. While we discuss here an actual binary indicator, any observable variable can fulfill this function. For instance, the user can monitor lamps indicating hard disk or network activity to decide whether a process has proceeded as intended. This variable may or may not be correlated with the actual process.

When signal detection is aided by a binary indicator, users should adopt different thresholds according to the information from the indicator. For instance, when the indicator notifies the user about a possible problem, the threshold should be lower (because with a valid indication system the prior probability of a problem is higher with such an indication) than when the system indicates that there is no